DOCUMENT RESUME

ED 039 241 TE 001 837

AUTHOR Clarke, Grace Allison
TITLE Interpreting the Penciled Scrawl: A Problem in Teacher Theme Evaluation.
PUB DATE May 69
NOTE 53p.; M.A. in Teaching Thesis, University of Chicago

EDRS PRICE EDRS Price MF-$0.25 HC-$2.75
DESCRIPTORS Academic Performance, *Composition (Literary), Composition Skills (Literary), English Instruction, Evaluation Criteria, *Evaluation Methods, Evaluation Techniques, Expository Writing, Secondary Education, Student Motivation, *Student Reaction, Success Factors, *Teacher Influence, Writing Skills

ABSTRACT

Examining student reactions to teachers' comments on themes, this study aimed to determine the effects of various combinations of reinforcement upon three dependent variables: (1) student perception of comments as positive or negative, (2) student satisfaction with comments, and (3) student confidence in writing ability. Six intermediate and advanced English classes (141 eleventh grade students in all) were divided randomly into nine treatment groups. Two cooperating teachers assigned an argumentative essay, due in 4 days, which then received one of nine comment treatments--composed by the experimenter--but no letter grade. Students were allowed 5 minutes to read these comments before the essays were collected and questionnaires were distributed to be filled out at leisure. Results showed that (1) the number of comments produced little effect; (2) purely negative comments produced lower scores in reinforcement, satisfaction, and confidence than completely positive comments produced; and (3) a mixture of criticism and praise, with praise dominating, produced the most satisfied and confident writers. (The writing assignment and questionnaire are appended.) (JM)

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

ED039241

THE UNIVERSITY OF CHICAGO

THE GRADUATE SCHOOL OF EDUCATION

INTERPRETING THE PENCILED SCRAWL: A PROBLEM IN TEACHER THEME EVALUATION

A PAPER SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF ARTS IN TEACHING

BY

GRACE ALLISON CLARKE

CHICAGO, ILLINOIS

MAY, 1969

TE001827

I INTRODUCTION

In 1924, Rollo Lyman had little praise for the comments written by teachers on student themes. He condemned typical methods of appraisal as mechanical and blamed them for making students hate to write. "The English composition class," he noted, "works too often under the unhealthy psychosis of fear and disgust when carrying through a written assignment."[1] Twenty-five years later, S. I. Hayakawa had the same complaint:

> Most freshman themes are written . . . in the knowledge that the incidental and even unconscious mechanics of the communicative act are going to be much more closely attended to than content. No wonder then, that the average freshman writes uneasily and self-consciously.[2]

The idea that written teacher comments can cause student frustration and discouragement is not new. Unfortunately, age does not guarantee an idea adequate analysis and validation. As late as 1963, Henry Meckel could write, "There appear . . . to have been no actual investigations of student attitudes and responses to teacher comments on papers."[3]

The aim of this study is to discover possible determinants of student attitudes toward written teacher comments. An underlying assumption is that encouragement and confidence constitute a better atmosphere for writing than "fear and disgust."

II REVIEW OF RESEARCH

Deficiencies of Existing Studies

There is no prior research on the question of student attitudes toward teacher comments themselves. Of the few studies examining the effect of written teacher comments on students, most deal solely with the effect of comments on performance. Only two studies consider both performance and, in addition, attitudes toward writing or English class.

It is most efficient to build on existing knowledge, and previous research on teacher comments could have proved helpful in constructing a research design for the present study. For example, before measuring student attitudes toward comments, one must identify the variables to which students might react. Any of the existing studies could be expected to include analysis of the variables differentiating one comment from another.

Unfortunately, most of the existing research is weakest in its analysis of variables. Relevant variables are often either unidentified or unclearly defined. The comment treatments studied by one researcher often differ radically and unsystematically from those studied by another. As a result, a tally of which research supports or rejects a particular type of comment is impossible. To make matters worse, the factor of correction or revision is an important--if disguised--variable in numerous instances. The most existing studies can do is indicate, by their deficiencies, questions which might be raised with profit.

Burton and Arnold (1963)

In The Effects of Frequency of Writing and Intensity of Teacher Evaluation Upon High School Students' Performance in Written Composition, Dwight Burton and Lois Arnold study the effect of two kinds of written teacher evaluation, moderate and intensive. Unfortunately, the authors' description of these independent variables is sufficiently vague and contradictory to make one suspect that the comment treatments were not uniform. According to the operational definition, moderate evaluation is "that kind of marking in which the teacher selects only an occasional paper to grade or corrects only those errors pertaining to skills which the students are studying at a particular time."[4] In the "Description of Treatments," however, one reads that moderate evaluation "concentrated one time on sentence structure, another on logic or organization."[5] One is left wondering whether or not all papers written by the moderate evaluation group received comment.

A similar problem of variable identification arises concerning the intensive evaluation groups. According to the operational definition, "intensive evaluation involves the marking of every error and the writing of detailed comments on each composition."[6] The "Description of Treatments," however, suggests that all types of errors did not receive equal attention. To quote this section: "Teachers were careful to mark every error

in mechanics, sentence structure, and usage and to write detailed comments on general effectiveness."[7] It is possible that in the intensive evaluation group, the emphasis was heavily on errors in form and mechanics to the possible neglect of shortcomings in content. It is not at all clear whether this emphasis was also present in the comment treatments of the moderate evaluation group. If not, the emphasis on form would have been a concealed, and possibly significant, independent variable.

Furthermore, the intensive evaluation treatment involved not only the number and kind of comments written on each paper, but also the additional variable of revision and rewriting. Students undergoing intensive evaluation were expected to revise and rewrite their papers and occasionally did this work outside class. Students in moderate evaluation groups neither revised nor rewrote nor worked on assigned essays outside class.

A researcher profiting from the mistakes of Burton and Arnold would clearly define his independent variables, being careful to identify and control such factors as revision and emphasis on form.

Buxton (1958)

In the work of Earl Buxton, as in that of Burton and Arnold, revision is tied to intensity of evaluation as a hidden variable. Experimental Group A received no grades and no marginal comments; they were not required to revise their essays.

The papers of Experimental Group B, on the other hand, received both grades and marginal comments; these students also revised and discussed their papers in class time.

Buxton, however, does take into account a variable totally ignored by Burton and Arnold: positive reinforcement. Both of Buxton's experimental treatments involved a paragraph of comment at the end of each paper. In Group A, the emphasis was on praise; each essay was praised as much as possible, and one or two ways in which the paper might be improved were pointed out. In Group B, no special effort at praise was made in the final paragraph of comment.[9]

It is worth noting that the number of instances of praise per paper was not controlled. Tight control of the independent variable of positive reinforcement would be a wise precaution for future experimenters.

Page (1958)

Positive reinforcement is treated in a similarly unsystematic way in the classic study by Ellis Page. Page used two experimental groups: a specified comment group and a free comment group. All his specified comments were designed to be encouraging; no similar restriction was placed on the free comments. The presence of either type of comment resulted in statistically significant improvements in objective test scores. Interestingly enough, there was no significant difference between the test performance of students in the two experimental groups.[10] One

might expect students to respond even more favorably to individual comments than to generalized ones; on the other hand, some free comments might have been less encouraging than the specified ones.

Furthermore, there is no indication that the free comments were comparable in length with the specified comments, which were about four or five words apiece. A free comment could be "whatever comment the teacher might feel it desirable to make."[11] Teachers were instructed: "Write anything that occurs to you in the circumstances. There is not any 'right' or 'wrong' comment for this study."[12]

If students equate length of comment with negativity of comment, comment length should be considered, in the future, as an important independent variable. In addition, positive and negative reinforcement should be controlled for both free and specified comments.

Sweet (1966)

In a study heavily indebted to Page, Roger Sweet studies the effect of written teacher comments on both student performance and student attitude. As might be expected, Sweet has the same possible sources of error as Page: he controls neither reinforcement nor length of comment.

Like Page, Sweet is concerned with the effects of comments written on tests "not of the long essay type,"[13] not

student compositions. Like Page, he used two experimental groups, a specified comment group and a free comment group. Unlike Page, however, Sweet distinguishes between short-term and long-run results. Since Page does not specify even the average time lapse between the two objective tests in his study, it is impossible to compare his results with Sweet's. Sweet concludes that both free and specified comments had little if any short-term effect on test performance. Over a longer period of time, however, free comments significantly improved both performance and attitudes.[14]

Sweet is concerned with attitude toward English class in general, not toward written comments in particular. His instrument is an inventory based on C. E. Osgood's semantic differential. All twenty-one items in this inventory are school-related activities; only six, however, are related to the English classroom in particular. Six questions provide scanty evidence on which to base conclusions. Furthermore, the items pertaining to the English classroom are: class discussions, oral reports, reading assignments, written assignments, library books, and length of the class period.[15] None of the English items measures attitude toward written English tests. This is a serious omission, since comments written on tests could be most reasonably expected to effect attitude toward tests.

Sweet's experience suggests the experimenter beware of

inadequate instruments. His results also indicate that the distinction between short-term and long-run can be very fruitful.

Taylor and Hoedt (1966)

Taylor and Hoedt examine, more systematically than any other experimenters, the variables of positive and negative reinforcement. The papers of Experimental Group A were praised regardless of actual merit; the best parts were circled in red; all errors were overlooked. Group B received heavy negative reinforcement; the errors on their papers were circled, and the best parts were overlooked.

Attitude in this study was assessed by a rather unsystematic variety of measures, including an attitude check list, anecdotal records of overt behavior, and the number of wrinkled torn papers (the assumption here is that such papers are an indication of anger and frustration).[16] Like that of Sweet, Taylor and Hoedt's study points to the difficulty of perfecting a good instrument for measuring attitudes.

III STATEMENT OF THE PROBLEM

This study focuses on two independent variables implicit--if frequently neither stated nor clearly defined--in the existing research: number and sign of written teacher comments. The aim is to determine the effects of various combinations of reinforcement upon three dependent variables: student perception of comments

as positive or negative, student satisfaction with comments, and student confidence in writing ability. Having done this, one can tentatively conclude which of a number of comment treatments are most conducive to student satisfaction and confidence. Because of the involved nature of the treatment groups, discussion of the hypotheses will be deferred until the necessary groundwork has been laid.

IV CONCEPTUALIZATION

Given the dearth of research, there can be no rigid model of how students react to the messages scribbled on their themes. One can safely assume, however, that this reaction is a composite. It seems best to begin, then, with a sketch of steps which might possibly contribute to the final reaction:

Step 1. Student decision on whether or not he understands the teacher's comments.

Step 2. Student interpretation of comments as primarily positive or primarily negative reinforcement.

Step 3. Student evaluation of the teacher as a reader of his paper: does the teacher understand the paper?

Step 4. Student evaluation of the teacher as a writer of comments: do the teacher's comments reflect biases along the dimensions of

a. content/expression?
b. positive/negative reinforcement?
c. number of comments?

Step 5. Student decision to accept or reject the teacher's biases.

Step 6. Student decision to agree or disagree with most of the comments written on his paper.

Step 7. Maintenance or adjustment by the student of his belief in the worth of his paper.

Step 8. Maintenance or adjustment by the student of his belief in his general writing ability.

One can reduce this scheme, with its eight sequential steps, to three distinct yet interrelated factors:

1. Student interpretation of comments as either positive or negative reinforcement. (Step 2.)

2. Student satisfaction with the comments. (Steps 1, 3, 4, 5, & 6; the student is assumed satisfied if he understands the comments and if his biases and judgments agree with those of the teacher.)

3. The effect of the comments on the student's confidence in his writing. (Steps 7 & 8.)

Factors 1, 2, and 3 were the basis for the construction of the experiment's attitude questionnaire. Note that while Factor 1 and Factor 2 are related, they are conceptually distinct. If, for whatever reason, the student is sufficiently dependent upon his teacher's judgment, the most negative of comments will in no way affect his "satisfaction" with what the teacher has written. Conversely, if the student has a high N_{Ach} and relatively low dependence, extensive negative reinforcement may lead to a general, and possibly ungrounded, dissatisfaction with the comments. Similarly, for those students with an inflated idea of their own abilities, negative reinforcement

may have little effect on their own estimation of their writing.

V RESEARCH DESIGN

Independent Variables

The two independent variables are number and sign of written teacher comments. These two elements were selected because of contemporary educational practice. Current thought equates adequacy of written evaluation with extensiveness of comment; the better the teacher, the more detailed the comment. On all except the very best papers, comments more often take the form of corrections than words of praise. Therefore, the more extensive the comment, the more numerous the instances of negative reinforcement.

Students, unlike educators, may interpret detailed comments as condemnations of their compositions, rather than as indicators of teacher concern. Such an interpretation may cause dissatisfaction with the comments and, at the extreme, complete rejection of them as meaningful statements. More importantly, detailed comments may be interpreted by the student to mean that his writing is a dismal failure and will always be so. Few of us elect to do those things at which we are destined to fail. The importance of discovering exactly how students do interpret lengthy and/or negative comments is obvious.

Dependent Variables

The dependent variables are the three factors outlined under "Conceptualization": student interpretation of the comments as either positive or negative reinforcement; student satisfaction with the comments; and student confidence in his writing.

Research Design

The research design sets up eight experimental groups and one control group:

Instances of Negative Reinforcement \ Instances of Positive Reinforcement	0	2-3	4-5
0	O	P_1	P_2
2-3	N_1	X_1	X_3
4-5	N_2	X_2	X_4

Explanation of Treatment Groups

Group O: zero instances of negative and positive reinforcement. The results of this group will be compared with data from the other groups to discover pretreatment levels of the independent variables.

Group N_1: two or three instances of negative reinforcement per page.

Group N_2: four or five instances of negative reinforcement per page.

Group P_1: two or three instances of positive reinforcement per page.

Group P_2: four or five instances of positive reinforcement per page.

Each N or P group measures the effect of one level of positive or negative reinforcement. It would be impossible to measure any interaction effects if the effect of each level of each type of reinforcement were unknown.

Group X_1: two or three instances each of positive and negative reinforcement per page.

Group X_2: two or three instances of positive reinforcement and four or five instances of negative reinforcement per page.

Group X_3: four or five instances of positive reinforcement and two or three instances of negative reinforcement per page.

Group X_4: four or five instances each of positive and negative reinforcement.

In all cases, an instance of positive or negative reinforcement is defined as one negative or positive free comment written in the margin of the student's paper. It was decided to use free comments in order to approximate actual class situations.

VI SAMPLE

The sample consisted of 141 eleventh grade students in six English classes. All subjects were drawn from a four year high school in an upper-middle to middle-middle class suburb of Chicago. There were four ability tracks at the school: Basic, Intermediate, Advanced, and Honors; subjects were taken only from the two middle tracks. Of the total number of subjects, 50 were enrolled in Intermediate English III and 91 in Advanced English III. All subjects were the students of one of two young female teachers. Each English teacher taught three classes.

VII METHODOLOGY

Treatment Groups

Treatment groups were formed by random stratified distribution. Each student was first ranked within his teacher's Advanced or Intermediate classes according to his most recent

quarterly grade in English. He was then randomly assigned to a group.

The Writing Assignment

On May 1, 1969, the cooperating teachers distributed instruction sheets for an argumentative essay due four days later. (See Appendix 1.) No additional directions for the assignment were given. All students wrote on the same topic in the same mode of discourse.

The assignment was designed to discourage purely or highly affective themes--and, in fact, none were submitted. This control was important since the affective content of a student's paper may influence his attitude toward comment treatments.

Writing Conditions

The writing conditions were as natural as possible. Since compositions were usually written outside class, students wrote their argumentative essays at home, free from imposed time limitations. This was an important experimental condition. If students are to write papers indicative of their abilities, they need time. In his "Suggested Methods of Research," Richard Braddock suggests that if high school students are not given a seventy to ninety minute writing period, "their sentence structure and mechanics will be produced under artificial circumstances."[17] In such situations, students might reject teachers' comments as totally irrelevant to their usual writing performance.

Comment Treatments

Each paper received one of the nine comment treatments. Since the experiment was designed to measure student reactions to comments only, no letter grade appeared on any paper. All comments were composed by the experimenter, but to normalize conditions, each teacher copied the comments onto the papers of her own students.

Administering the Questionnaire

Papers were returned during a regular class period. As soon as students received their essays, they were directed as follows:

> For the next five minutes, look over whatever comments happen to be on your papers. It will be important for you to be familiar with them. Your grades are not on your papers, but they are recorded in my grade book. I will tell you your grades later in the period.

Five minutes later, the essays were collected; students could not reread the comments while answering the questionnaires. If the students had kept their papers, their responses might have indicated more objectivity than they actually felt. Our purpose was to use the questionnaire as a measuring instrument, not as a study guide for the comments.

After the papers were collected, the questionnaires[18] (see Appendix 2) were distributed, and each teacher read aloud the explanatory first page to her class. Students were then free to respond to the questionnaire at their own speeds.

Statistical Analysis

After assigning a number to every possible answer,[19] we calculated the arithmetic mean of each group's response to each question ($\bar{X}_{ij}$). Using $\bar{X}_{ij}$ and the estimated variance (pooled) of all combinations of two means, we determined the significant differences among groups for each question.[20] A standard one-tail T-test measured significance at the .05 level. (See Appendix 3.)

Reinforcement, Satisfaction, and Confidence Scores

The next step was to discover the relative effects of the comment treatments upon the dependent variables. Inspection of $\bar{X}_{ij}$ shows that some comment treatments caused responses which are significantly high or low relative to others. The following Relativity Scale quantitizes these differences:

A score of six: a response significantly higher than all others.

A score of five: a response significantly higher than some but similar to others.

A score of four: a response significantly higher and lower than some but similar to others.

A score of three: a response similar to all others.

A score of two: a response significantly lower than some but similar to others.

A score of one: a response significantly lower than all others.

The raw group scores for relative satisfaction, confidence,

and positive reinforcement were found by summing group scores for the appropriate questions.[21] To discover the changes caused by the experimental comment treatments, we subtracted the relevant Group O score from each of the twenty-four composite experimental group scores. (Remember that the scores of Group O indicate pretreatment levels of the dependent variables.) These adjusted scores were the final measures of Satisfaction (S), Confidence (C), and Reinforcement (R).

VIII HYPOTHESES

If students have been so conditioned by grading procedures that they blindly equate detailed response with negative response, Group X_4 (the group with the highest number of total comments) will score consistently lower than all other groups; X_2 and X_3 will score the second lowest; X_1, P_2, and N_2 will score second highest; and P_1 and N_1 will score highest.

If each score of each mixed treatment group (X_1, X_2, X_3, and X_4) is not equal to the sum of the scores of that group's components, then there is interaction between the instances of positive and negative reinforcement in that group.

Briefly, the purpose of the experiment is to disprove the following null hypotheses:

1. There is no difference between pretreatment and post treatment levels of S, C, and R.

$0 = N_1 = N_2 = P_1 = P_2 = X_1 = X_2 = X_3 = X_4$

2. There is no difference in the effect of positive and negative reinforcement on S, C, and R.

$N_1 = P_1 \qquad N_2 = P_2$

3. There is no difference in the effect of different amounts of positive reinforcement on S, C, and R.

$P_1 = P_2$

4. There is no difference in the effect of different amounts of negative reinforcement on S, C, and R.

$N_1 = N_2$

5. There is no interaction between positive and negative reinforcement.

$X_1 = N_1 + P_1;\ X_2 = N_2 + P_1;\ X_3 = N_1 + P_2;\ X_4 = N_2 + P_2$

IX RESULTS

Omitted Questions

Since group responses to neither question four nor question six differed significantly, these questions were omitted in the final scoring. Question four measured students' confidence in the quality of their next paper; question six measured opinions on the relative importance of form and content to good writing. Of 141 students, 65 said they would be

"somewhat pleased with the quality of their next paper; 30, "very pleased"; and 25, "neither pleased nor displeased." Only 10 said they would be "somewhat displeased"; only 2, "very displeased." The 9 remaining students left the question blank. Of 141 students, 99 thought form and content equally important; 33 favored content over form; 9 favored form over content.

Number of Comments; Amount of Praise; Amount of Criticism

How did students respond to those questions which directly measured their reactions to the independent variables?

The mean response (3.0) of treatment group N_1 was equivalent to the optimal response, "just enough comments." Group N_1's response was significantly closer to this optimum than that of groups P_2, X_4, and N_2; it's response was similar, however, to that of all other groups. Comment treatments X_4 and N_2 were considered to contain significantly more than the optimum number of comments; treatment P_2, significantly fewer.

No comment treatment contained significantly more than the optimal amount of criticism; treatments P_1, P_2, and X_1, however, all contained significantly less. Group X_3 (mean response of 3.0) was significantly closer to this optimum than P_2, P_1, and X_1; the response of all other groups was similar to that of X_3.

The mean response (3.0) of treatment group X_4 was equivalent to the optimal response, "just enough praise."

Treatments N_1 and N_2 contained less than the optimal amount of praise; treatments P_2 and P_1, significantly more. All other groups could not be said to contain significantly more or less than "just enough praise."

Reinforcement, Satisfaction, and Confidence Scores

All Reinforcement, Satisfaction, and Confidence scores are reported below. In no case will differences of +1 or -1 be considered as meaningful.

	O	P_1	P_2	N_1	N_2	X_1	X_2	X_3	X_4
R	0	0	2	-2	-2	-2	-2	0	-2
C	0	7	10	-1	0	1	-1	10	6
S	0	7	7	3	-2	5	3	11	1

X CONCLUSIONS

Pretreatment Conditions

If nothing else, this study indicates that "fear and disgust" still are rampant in the English classroom. It is rather sobering to realize that the Satisfaction score for Group O was lower than that of any other group but Group N_2. The Confidence score for Group O was lower than that of

four experimental groups; it was higher than none. There is one bright spot, however: most students seem to recover, at least partially, from the crushing effects of written teacher comments. After all, more than two-thirds thought they would be either "somewhat pleased" or "very pleased" with the results of their next paper.

Form vs Content

The neglect of content emphasized so heavily by Lyman and Hayakawa does not seem a relevant factor for the classes in this sample. As mentioned earlier, most students thought form and content equally important to good writing. Furthermore, only 21 out of 138 thought their teachers valued form more than content. Of the rest, all but 7 felt their teachers considered form and content equally important. If there was not an undue emphasis on form, what did cause dissatisfaction with written teacher comments? Or, to turn the question around, what sort of comment treatments lead to the highest satisfaction?

Number of Comments

Sheer number of comments does not determine how a comment treatment will be received. Neither R, C, nor S was dependent upon number of comments. Those treatment groups with the fewest comments--P_1, N_1, X_1, P_2, and N_2--did not produce the highest R's, C's, or S's. Nor did those treatment

groups with the most comments--X_4, X_2, and X_3--score correspondingly low.

Experimental conditions required students to study their comments attentively. This suggests, at the very least, that when students are forced to read the comments on their papers, they are not biased by number.

Positive and Negative Reinforcement

The results showed that there was a difference in effect between positive and negative reinforcement. The purely negative comment treatments, N_1 and N_2, generally produced scores lower than the initial (pretreatment) R, C, and S. As might be expected, the purely positive treatments, P_1 and P_2, generally produced scores higher than those of Group O. There were only three exceptions to this pattern: treatment N_1 caused an increase in S; treatment N_2 caused no change in C; treatment P_1 caused no change in R.

Levels of Positive Reinforcement

In two instances out of three, Group P_2 scored higher than Group P_1. In the case of S, however, scores for Group P_1 and Group P_2 were equivalent. Higher levels of purely positive reinforcement produced scores equal to, or higher than, those caused by lower levels.

Levels of Negative Reinforcement

In the case of R and C, the scores of Group N_1 and Group N_2 were equivalent. The S of Group N_1, however, was larger by five than the S of Group N_2. This is a large difference in a total range of thirteen points. It is important to remember here that R and C together are a composite of only five questions, while S is a composite of eight. In this light, it seems reasonable to consider the difference between S's an important one. Whereas Treatment N_1 raised S from the pretreatment level, Treatment N_2 decreased it, by roughly the same amount. This suggests that once criticism exceeds what is perceived as a "normal" level, satisfaction with the comments will decrease. Presumably, the S measure, with its multiple questions, caught a difference undetected by the single question (#13) which measured response to the amount of criticism. As noted earlier, no comment treatment contained more than the optimal amount of criticism--although both N_1 and N_2 contained less than the optimal amount of praise. Perhaps students perceive negative loading of comments as "too little praise," not "too much criticism."

Interaction Effects

Of the four mixed treatment groups, only Group X_3 showed no evidence of interaction between instances of positive and negative reinforcement. In the cases of X_1, X_2,

and X_4, there were large differences between the score of the X group and the combined scores of its components. The only two exceptions to this pattern were the R's for Group X_1 and Group X_2. This deviation may be explained by the severely limited range (/2/) between the R's of Group N_1, Group P_1, and Group P_2. It seems safe to conclude that interaction exists in all mixed groups except Group X_3.

What is the significance of the lack of interaction in the case of Group X_3? One of the components of Group X_3 is P_2, which is consistently a very high scorer. Group X_3 would have had to score high indeed to surpass the combined score of its components. Those who think praise a panacea should find it remarkable that a treatment containing negative reinforcement produced scores as high as those of the heavily positive treatment P_2. Remember: treatment P_2 was considered to contain more than the optimal amount of praise and less than the optimal amount of criticism. There is such a thing as "too much praise" and "too little criticism."

<u>Recommendations</u>

Careful balancing of negative comments with positive is not an adequate solution to the problem of "fear and disgust." The C of Group X_1 is in the lower half of all Confidence scores, while the S of Group X_4 is in the lower half of all S's. Students reacted with the greatest

confidence and the most satisfaction to treatments X_3, P_1, and P_2. These were the same groups which they perceived as relatively positive--and which were, in fact, positively loaded.

The implications seem clear. An occasional positive comment is not enough. Teachers should be sure they find more to praise than to criticize. We do not suggest that all criticism be omitted. In fact, the indications are that criticism, well tempered with praise, produces the most satisfied and confident writers of all.

APPENDIX 1: THE WRITING ASSIGNMENT

The Results of Violence on Television

Violence on television is the cause of increased violence today in the streets and on college campuses. The average evening T.V. show is a class in crime: blackmail, murder, robbery, riots, and drugs. Young people often get into trouble merely for copying behavior which adults allow to be shown on the screen. Children grow up glued to the T.V. set, learning as they watch that violent action is the answer to their problem--whether the problem is poverty, discrimination, neglect, or simple boredom. Even the evening news broadcast

features only blood and gore. The situation has grown so bad that there now remains just one solution. We, as parents, must totally refuse to allow our children to watch television. At the worst, our youngsters will read more good books and play more baseball. If they get bored, they can always gather around the radio. After all, that's what we did when we were young, and we're still a pretty peaceful bunch, aren't we?

Write a well-organized, well-supported essay in which you comment on this paragraph. You may find the following questions helpful:

To what extent do <u>you</u> agree or disagree with this statement? Remember that this statement has many parts; consider each one of them carefully. <u>Support your stand with examples</u>. Be sure to refer to specific events and/or specific T.V. programs. Is this writer objective or biased? Point out specific words or phrases which support your answer. What sort of person would be likely to agree with this man? To disagree with him? Why?

APPENDIX 2: QUESTIONNAIRES, VERSIONS A AND B

Explanatory Page Attached to Questionnaire, Versions A and B

The following questionnaire is part of a study designed to find out what students think if the comments which teachers write on their compositions. We want to know what you think.

Please be completely honest in your response. Your answers will be entirely anonymous; there is no way of telling which questionnaire belongs to which student. What you say cannot possibly get you (or your teacher) into trouble.

Remember: you are playing an important part in a piece of educational research. Your opinions can help teachers teach you better.

Questionnaire Version A, Administered to Group O

1. On the basis of the comments usually written on my English papers, I think my teacher would agree that:

   a. My English papers are very poor.
   b. Despite some serious errors, my English papers have their good points.
   c. There is nothing either particularly good or particularly bad about my English papers.
   d. Despite some minor errors, my English papers have their good points.
   e. My English papers are very good.

2. I feel that I usually understand:

   a. all of the comments written on my English papers.
   b. most of the comments written on my English papers.
   c. some of the comments written on my English papers.
   d. few of the comments written on my English papers.
   e. none of the comments written on my English papers.

3. I usually agree with:

   a. none of the comments written on my English papers.
   b. few of the comments written on my English papers.
   c. some of the comments written on my English papers.
   d. most of the comments written on my English papers.
   e. all of the comments written on my English papers.

4. I personally would agree with the following statement:

   a. "It's not <u>what</u> you say--it's <u>how</u> you say it."
   b. "What a person has to say is more important than how he says it."
   c. "In good writing, what a person says and how he says it are equally important."

5. I usually feel that:

   a. all of the mistakes in my English papers are minor.
   b. most of the mistakes in my English papers are minor.
   c. some of the mistakes in my English papers are minor.
   d. few of the mistakes in my English papers are minor.
   e. none of the mistakes in my English papers are minor.

6. I feel that when I finish writing my next English paper, I will be:

   a. very displeased with its quality.
   b. somewhat displeased with its quality.
   c. neither pleased nor displeased with its quality.
   d. somewhat pleased with its quality.
   e. very pleased with its quality.

7. I feel that I write:

   a. much better than most of the students in my class.
   b. a little better than most of the students in my class.
   c. as well as most of the students in my class.
   d. a little worse than most of the students in my class.
   e. much worse than most of the students in my class.

8. The following statement applies to the number of comments usually written on my English papers:

   a. there are far too many comments.
   b. there are too many comments.
   c. there are just enough comments.
   d. there are too few comments.
   e. there are far too few comments.

9. Usually, after reading the comments on one of my English papers, I personally think that the paper:

   a. is very poor.
   b. despite some serious errors, has its good points.
   c. is neither particularly good nor particularly bad.
   d. despite some minor errors, has its good points.
   e. is very good.

10. I feel that the amount of praise usually contained in the comments on my paper:

    a. is far too little.
    b. is too little.
    c. is just enough.
    d. is too much.
    e. is far too much.

11. I usually feel that:

    a. all of the mistakes in my English papers are major.
    b. most of the mistakes in my English papers are major.
    c. some of the mistakes in my English papers are major.
    d. few of the mistakes in my English papers are major.
    e. none of the mistakes in my English papers are major.

12. Judging from the comments written on my English papers, my teacher usually understands my papers:

    a. completely.
    b. almost completely.
    c. somewhat.
    d. not at all.

13. I feel that the amount of criticism usually contained in the comments on my English papers:

    a. is far too much.
    b. is too much.
    c. is just enough.
    d. is too little.
    e. is far too little.

14. Judging from the comments usually written on my English papers, my teacher would most agree with the following statement:

    a. "It's not <u>what</u> you say--it's <u>how</u> you say it."
    b. "What a person has to say is more important than how he says it."
    c. "In good writing, what a person says and how he says it are equally important."

<u>Questionnaire Version B, Administered to Experimental Groups</u>

1. On the basis of the comments written on my paper, I think my teacher would agree that:

    a. This paper is very poor.
    b. Despite some serious errors, this paper has its good points.
    c. There is nothing either particularly good or particularly bad about this paper.
    d. Despite some minor errors, this paper has its good points.
    e. This paper is very good.

2. I feel that I understand:

a. all of the comments written on this paper.
b. most of the comments written on this paper.
c. some of the comments written on this paper.
d. few of the comments written on this paper.
e. none of the comments written on this paper.

3. I agree with:

a. none of the comments written on this paper.
b. few of the comments written on this paper.
c. some of the comments written on this paper.
d. most of the comments written on this paper.
e. all of the comments written on this paper.

4. I personally would agree with the following statement:

a. "It's not <u>what</u> you say--it's <u>how</u> you say it."
b. "What a person has to say is more important than how he says it."
c. "In good writing, what a person says and how he says it are equally important."

5. I feel that:

a. all of the mistakes in this paper are minor.
b. most of the mistakes in this paper are minor.
c. some of the mistakes in this paper are minor.
d. few of the mistakes in this paper are minor.
e. none of the mistakes in this paper are minor.

6. I feel that when I finish writing my next paper, I will be:

a. very displeased with its quality.
b. somewhat displeased with its quality.
c. neither pleased nor displeased with its quality.
d. somewhat pleased with its quality.
e. very pleased with its quality.

7. I feel that I write:

   a. much better than most of the students in my class.
   b. a little better than most of the students in my class.
   c. as well as most of the students in my class.
   d. a little worse than most of the students in my class.
   e. much worse than most of the students in my class.

8. The following statement applies to the number of comments on my paper:

   a. There are far too many comments.
   b. There are too many comments.
   c. There are just enough comments.
   d. There are too few comments.
   e. There are far too few comments.

9. After reading the comments on my paper, I personally think that this paper:

   a. is very poor.
   b. despite some serious errors, has its good points.
   c. is neither particularly good nor particularly bad.
   d. despite some minor errors, has its good points.
   e. is very good.

10. I feel that the amount of praise contained in the comments on my paper:

    a. was far too little.
    b. was too little.
    c. was just enough.
    d. was too much.
    e. was far too much.

11. I feel that:

    a. all of the mistakes in this paper are major.
    b. most of the mistakes in this paper are major.
    c. some of the mistakes in this paper are major.
    d. few of the mistakes in this paper are major.
    e. none of the mistakes in this paper are major.

12. Judging from the comments written on my paper, my teacher understood my paper:

    a. completely.
    b. almost completely.
    c. somewhat.
    d. not at all.

13. I feel that the amount of criticism contained in the comments on this paper:

    a. was far too much.
    b. was too much.
    c. was just enough.
    d. was too little.
    e. was far too little.

14. Judging from the comments on my paper, my teacher would most agree with the following statement:

    a. "It's not <u>what</u> you say--it's <u>how</u> you say it."
    b. "What a person has to say is more important than how he says it."
    c. "In good writing, what a person says and how he says it are equally important."

APPENDIX 3: TABLES OF SIGNIFICANCE

QUESTION 1*

	X_4	X_3	P_2	X_2	X_1	P_1	N_2	N_1	O
X_4		S	S	N	N	S	N	N	N
X_3			S	S	S	S	S	S	N
P_2				S	S	S	S	S	S
X_2					N	S	N	N	N
X_1						S	N	N	N
P_1							S	S	S
N_2								N	S
N_1									S

* In this table, as in all those to follow, "S" means significant at .05, while "N" means not significant at .05.

QUESTION 2

	X_4	X_3	P_2	X_2	X_1	P_1	N_2	N_1	O
X_4		S	S	N	N	S	N	N	N
X_3			N	S	N	N	S	N	S
P_2				N	N	N	N	N	S
X_2					N	S	N	N	N
X_1						N	N	N	N
P_1							S	S	S
N_2								N	N
N_1									S

QUESTION 3

	X_4	X_3	P_2	X_2	X_1	P_1	N_2	N_1	O
X_4		S	S	N	N	S	N	N	N
X_3			N	S	N	N	S	S	S
P_2				S	N	N	S	S	S
X_2					N	S	N	N	N
X_1						N	N	N	N
P_1							S	N	S
N_2								N	N
N_1									N

QUESTION 5

	X_4	X_3	P_2	X_2	X_1	P_1	N_2	N_1	O
X_4		N	S	N	N	N	S	N	S
X_3			N	N	N	N	S	N	S
P_2				S	S	N	S	S	S
X_2					N	S	N	N	N
X_1						S	N	N	N
P_1							S	S	S
N_2								N	N
N_1									N

QUESTION 7

	X_4	X_3	P_2	X_2	X_1	P_1	N_2	N_1	O
X_4		N	N	N	N	N	N	N	N
X_3			N	S	S	S	N	S	N
P_2				N	N	N	N	N	N
X_2					N	N	N	N	N
X_1						N	N	N	N
P_1							N	N	N
N_2								N	N
N_1									N

QUESTION 8

	X_4	X_3	P_2	X_2	X_1	P_1	N_2	N_1	O
X_4		N	N	N	N	N	N	S	N
X_3			N	N	N	N	N	N	N
P_2				N	N	N	N	S	N
X_2					N	N	N	N	N
X_1						N	N	N	N
P_1							N	N	N
N_2								S	N
N_1									N
O									

QUESTION 9

	X_4	X_3	P_2	X_2	X_1	P_1	N_2	N_1	O
X_4		S	S	N	N	N	S	N	N
X_3			S	S	N	N	S	S	S
P_2				S	S	S	S	S	S
X_2					N	S	N	N	N
X_1						N	S	N	N
P_1							S	S	S
N_2								N	N
N_1									N

QUESTION 10

	X_4	X_3	P_2	X_2	X_1	P_1	N_2	N_1	O
X_4		N	S	N	N	S	S	S	N
X_3			N	N	N	N	S	S	N
P_2				S	S	N	N	S	S
X_2					N	S	S	S	N
X_1						S	S	S	N
P_1							S	S	S
N_2								N	S
N_1									S

QUESTION 11

	X_4	X_3	P_2	X_2	X_1	P_1	N_2	N_1	O
X_4		N	N	N	N	S	S	S	N
X_3			N	S	S	N	S	S	S
P_2				S	S	N	S	S	S
X_2					N	S	N	N	N
X_1						S	N	N	N
P_1							S	S	S
N_2								N	N
N_1									N

QUESTION 12

	X_4	X_3	P_2	X_2	X_1	P_1	N_2	N_1	O
X_4		S	S	N	N	S	N	S	N
X_3			N	S	S	S	S	S	S
P_2				S	S	N	S	S	S
X_2					N	S	N	S	N
X_1						S	N	S	N
P_1							S	S	S
N_2								N	N
N_1									S

QUESTION 13

	X_4	X_3	P_2	X_2	X_1	P_1	N_2	N_1	O
X_4		N	S	N	N	S	N	N	N
X_3			S	N	S	S	N	N	N
P_2				S	S	N	S	S	S
X_2					N	S	N	N	N
X_1						S	N	N	N
P_1							S	S	S
N_2								N	N
N_1									N

QUESTION 14

	X_4	X_3	P_2	X_2	X_1	P_1	N_2	N_1	O
X_4		N	N	N	N	N	N	S	N
X_3			N	N	N	N	N	S	N
P_2				N	N	N	N	N	N
X_2					N	N	N	S	N
X_1						N	N	S	N
P_1							N	N	N
N_2								N	N
N_1									N

QUESTION 15*

	X_4	X_3	P_2	X_2	X_1	P_1	N_2	N_1	O
X_4		N	N	N	N	N	N	N	N
X_3			N	N	N	S	N	N	S
P_2				N	N	S	N	N	S
X_2					N	N	N	N	N
X_1						N	N	S	N
P_1							S	S	N
N_2								N	S
N_1									S

*Question 15 is a comparison of responses to Questions 1 and 9.

QUESTION 16*

	X_4	X_3	P_2	X_2	X_1	P_1	N_2	N_1	O
X_4		N	N	N	N	N	N	N	N
X_3			N	N	N	S	N	N	S
P_2				N	N	S	N	N	S
X_2					N	N	N	N	N
X_1						N	N	S	N
P_1							S	S	N
N_2								N	S
N_1									S

* Question 16 is a comparison of responses to Questions 4 and 14.

FOOTNOTES

[1] Rollo L. Lyman, Summary of Investigations Relating to Grammar, Language, and Composition (Supplementary Educational Monographs; Chicago: University of Chicago, 1929), p. 188.

[2] S. I. Hayakawa, "Linguistic Science and Teaching Composition," A Guide For Evaluating Student Composition, ed. Sister M. Judine, I. H. M. (Champaign, Illinois: National Council of Teachers of English, 1965), pp. 1-2.

[3] Henry C. Meckel, "Research on Teaching Composition and Literature," Handbook of Research on Teaching (Chicago: Rand McNally and Co., 1963), p. 986.

[4] Dwight L. Burton and Lois V. Arnold, Effects of Frequency in Writing and Intensity of Teacher Evaluation Upon High School Students' Performance in Written Composition (Cooperative Research Project No. 1523; Tallahassee, Florida: Florida State University, 1963), p. 28.

[5] Ibid., p. 40.

[6] Ibid., p. 28; italics added.

[7] Ibid., p. 41.

[8] Ibid., p. 28.

[9] Richard Braddock, Richard Lloyd-Jones, and Lowell Schoer, "Summaries of Selected Research: The Buxton Study," Research in Written Composition (Champaign, Illinois: National Council of Teachers of English, 1963), pp. 58-70.

[10] Ellis Batten Page, "Teacher Comments and Student Performance: A Seventy-Four Classroom Experiment in School Motivation," The Journal of Educational Psychology, 49, p. 181.

[11] Ibid., p. 174.

[12] Ibid., p. 174.

[13] Roger C. Sweet, Educational Attainment and Attitudes Toward School As A Function of Feedback In The Form of Teachers' Written Comments (Unpublished Master's Thesis, Available Through ERIC; Madison: University of Wisconsin, November, 1966), p. 1.

[14] Ibid., p. 1.

[15] Unpublished attitude survey forwarded to author by Roger C. Sweet.

[16] Winnifred Taylor and Kenneth C. Hoedt, "The Effect of Praise Upon the Quality and Quantity of Creative Writing," Journal of Educational Research, LX (October, 1966), pp. 80-81.

[17] Braddock, Jones, and Schoer, "Suggested Methods of Research, " p. 9.

[18] Two versions of the questionnaire were used, one for Group 0 and another for the experimental groups. Since Group 0 papers received no comments, in the Group 0 questionnaire "usually" was inserted in all questions. Otherwise the questionnaires were identical.

[19] For all questions except 8, 10, and 13, the highest numerical equivalent was assigned to that response which indicated the greatest perceived positive reinforcement, the greatest satisfaction, or the greatest confidence. In the case of questions 8, 10, and 13, a response of "far too little" or "far too few" was assigned a value of 5; a respons of "too little" or "too few"

a value of 4; a response of "just enough" a value of 3; a response of "too many or "too much" a value of 2; and a response of "far too many" or "far too much," a value of 1.

[20] For questions 8, 10, and 13, the absolute values of the differences of the means from 3 were compared to determine significance.

[21] The raw score for satisfaction is the sum of the scores of questions 2, 3, 8, 10, 12, 13, 15, and 16. The raw score for confidence is the sum of the scores of questions 7, 9, 5, and 11. The raw score for positive reinforcement is equivalent to the score for question 1.

BIBLIOGRAPHY

Braddock, Richard; Lloyd-Jones, Richard; and Schoer, Lowell. Research in Written Composition. Champaign, Illinois: National Council of Teachers of English, 1963.

Burton, Dwight L. and Lois V. Arnold. Effects of Frequency in Writing and Intensity of Teacher Evaluation Upon High School Students' Performance in Written Composition. Tallahassee, Florida: Florida State University, 1963. Cooperative Research Project No. 1523.

Denby, Robert V. "NCTE/ERIC Report--Composition Evaluation," The English Journal (November, 1968), pp. 1215-1221.

Judine, Sister M., I. H. M. A Guide for Evaluating Student Composition: Readings and Suggestions for the Teachers of English in the Junior and Senior High School. Champaign, Illinois: National Council of Teachers of English, 1965.

Lyman, Rollo L. Summary of Investigations Relating to Grammar, Language, and Composition. Chicago, Illinois: University of Chicago, 1929. Supplementary Educational Monographs.

Meckel, Henry C. "Research on Teaching Composition and Literature," Handbook of Research on Teaching.

Edited by N. L. Gage. Chicago: Rand McNally and Company, 1963, pp. 966-1001.

Osgood, C.E.; Suci, G. J.; and Tannenbaum, P. H. The Measurement of Meaning. Urbana: University of Illinois Press, 1957.

Page, Ellis Batten. "Teacher Comments and Student Performance: A Seventy-Four Classroom Experiment in School Motivation." The Journal of Educational Psychology, 49 (1958), pp. 173-181.

Sweet, Roger C. Educational Attainment and Attitudes Toward School as a Function of Feedback in the Form of Teachers' Written Comments. Madison: University of Wisconsin, November, 1966.

_____. Unpublished attitude inventory.

Taylor, Winnifred and Kenneth C. Hoedt. "The Effect of Praise Upon the Quality and Quantity of Creative Writing." The Journal of Educational Research, LX (October, 1966), pp. 80-84.